Bacteriophage 2:1, 55–59; January/February/March 2012; © 2012 Landes Bioscience

Who went into phage research?

Hans-W. Ackermann

Department of Microbiology, Immunologie and Infectiologie; Faculty of Medicine; Laval University; Quebec, Canada

A total of 30,000 phage papers, books, or book chapters, published between 1965 and 2010, were analyzed for the ethnic origins of 14,429 first authors. Their names represent 40 linguistic domains or geographic areas and at least 70 languages. British and German names predominate. Results broadly concur with statistics on the frequency of publications by country and show the growing role of Third-World countries in phage research. Irish and Jewish scientists are prominent. Historical and societal factors appear to be very important elements in the advancement of science.

© 2012 Landes Bioscience. Do not distribute.

Keywords: bibliography, cultural theory, ethnic origin, family names, linguistics

Submitted: 10/31/11

Revised: 11/07/11

Accepted: 11/07/11

http://dx.doi.org/10.4161/bact.18680

Correspondence to: Hans-W. Ackermann; Email: ackermann@mcb.ulava.ca

Introduction

Bacteriophages are among the best-studied microbes in science. Bacteriophage scientists are in the fortunate position of having several literature collections at their disposal. In 1958 and 1967, H. Raettig in Germany published two phage bibliographies which covered phage research from its beginnings in 1915 to 1965 and totalized 11,358 references.[1-3] As a student of Raettig, I was duly impressed. Since then, the volume of phage literature has progressed enormously and can now be estimated at 45,000 publications, excluding conference abstracts, technical reports and patent applications. This constitutes a unique body of literature for analysis. I analyzed 30,000 phage publications, published between 1965 and 1930 worldwide, for the origins of 14,429 first authors. The body of 30,000 publications is based on my own extensive literature collection, those of the Bacteriophage Ecology Group (BEG)[4] and databases such as Current Contents and Scopus (Elsevier). It is now available on the Net (www.phage.ulaval.ca). For purposes of comparison, this analysis was extended to the phage bibliographies of Raettig[1,2] (unpublished).

The study of scientific publication patterns is called Bibliometrics and measures scientific output according to fields, country, language, funding and impact by citation frequency. It has provided many insights, for example into the dominant role of English as language of science[5] or links between funding and productivity. However, to my knowledge, it has never been attempted to sort out scientists by ethnic origins, although this relates directly to current theories on scientific advance, whether cultural, structural, or merely linked to the number of scientists.

The cultural theory proposes that scientific advancement is linked to societies that place a high value on scientific activity. It was formulated in 1938 by E.K. Merton who linked English puritanism and German pietism, thus ascetic Protestant varieties, to the scientific revolution in the 17th and 18th centuries.[6] In doing so, he overlooked Catholic scientists such as Galileo or Copernicus. The structural theory considers competition between universities and job opportunities as the most important factors, whereas the numerical theory emphasizes the number of available scientists.[7] Scientific research has long been the preserve of a few wealthy countries, primarily from Europe and North America. The economic emergence of Third-World countries, of instant communication via the Net, and international scholarships have created an entirely new situation. In this context, what can be said about virologists and phage scientists in particular?

Identifying Names

This study is limited to surnames of first (often only) authors of a given paper. This gives graduate students from Third-World countries an edge as they are commonly listed first in papers with multiple authors. By contrast, established scientists are generally listed last. It is rather futile to analyzse all names in all papers, as some publications, especially those on genome sequences and metagenomics, list 50 to 100 contributors. Similarly, I did not attempt to sort names according to the institutes of investigation. This would essentially result in quantifying productivity by country, hence a subject that has been covered repeatedly.[5,6,8,9] Also, publications are increasingly often the joint products of several laboratories. For example, I recently saw a paper by eight scientists from seven laboratories and four countries.[10]

Most surnames fall into four categories, those derived from given names (e.g., Jackson or Johnson), occupational (Fisher, Smith), topographic (Hill, Wood), or nicknames (White, Black).

In Europe, surnames originated in the 10th century and were often attributes of the nobility. They became gradually generalized, but words and names were spelled over centuries according to sound, generating countless variants. Until the advent of dictionaries, the English language lacked a system of spelling rules. A few peoples still have no surnames; this is why some Indian scientists have a single name only. Jews were required in many countries, often as late as the 18th century, to invent themselves family names. By contrast with the West, the Chinese adopted hereditary surnames as early as 5000 years ago.

The identification of names requires some knowledge of languages and a great deal of linguistics. This author is familiar with English, French, German, and Spanish, has learned and forgotten Farsi and Japanese, has recently delved into African languages (Fon, Kikongo, Yoruba), and has a keen interest in linguistics. Many languages, and hence names derived from them, have a characteristic morphology. For example, Finnish, Georgian, Japanese, or Scottish names are generally easy to identify. The contrary is true for names from Iran and Indonesia. The Net is an invaluable tool here. It offers excellent sites devoted to genealogy, for example "houseofnames.com," as well as lists of Basque, Breton, Irish, Korean, Scottish, Ukrainian, or Welsh names, and even long and detailed dictionaries of Jewish family names. I was also able to rely on the help of friends or colleagues from Algeria, India, Ukraine and Turkey. In difficult cases, the origin of a paper or the names of coauthors helped to identify a name.

Sources of Error

One major problem was the occurrence of identical names in groups of related languages, for example German-Dutch-Danish, Russian-Ukrainian-Polish, or Scottish-Irish-Welsh. Surnames were sometimes impossible to classify without consulting the original papers where they occurred or without establishing their preferential geographical distribution. Particular problems were that (a) Jews often adopted common German or English family names, (b) Chinese and Korean surnames are sometimes identical and must then be identified by the first names of their owners, and (c) Russian names are frequently of mixed origin, reflecting the expansionism of the Tsarist empire. For example, I knew an excellent Russian electron microscopist named Azizbekyan. "Aziz" is an Arabic first name, "Bek" or "Beg" is a Turkish title, and "Yan" is Armenian for son. Many Russian names are ultimately of Persian, Kyrgyz, Tatar, Finno-Ugrian, or Chechen origin.

It also happened that, upon immigration to the USA, many immigrants changed names. This could be through the good services of immigration officials at Ellis Island, New York, who, apparently, were legendary in the fabrication of English names. It seems that many names of Polish or Russian origin fared badly and were anglicized by error. Some immigrants, especially Germans, changed names deliberately, morphing from "Müller" to "Miller" or from "Fischer" to "Fisher." This suggests that the number of phage scientists of German origin is much higher than apparent. Another problem resides in the fact that there are only 200 Chinese family names. I checked their first names, but, even so, the number of Chinese scientists may be underreported here because of homonyms. Thus, the science of names is fairly inexact and I may have accumulated countless errors.

Findings

The analyzed 30,000 phage publications have over 14,500 first authors. Name lists are available on request. The very high number of authors is due to that fact junior scientists, many of which may write no more than one paper in their life, are given an edge in this analysis. The publications belong to 40 linguistic domains or geographic areas and at least 70 languages (**Tables 1** and **2**). The total number of the latter is difficult to determine because several African names were not identified and no detailed identification of Indian languages was attempted. The vast majority of names (75%) are derived from European languages. The frequency of names per language domain roughly parallels the number of scientific publications by country.[6,8]

The names are silent as to nationality, residence, past, career, religion, or workplace of their owners or the ancestors of these. Gender is evident in a number of Slavic or Icelandic names, but there is no provision for women who changed their names when they married. The names indicate ancestry or no more. For example, a French name may belong to citizens of France, Belgium, French Canada, Haiti, or any country where this citizen or his ancestor decided to live. Such an ancestor might have immigrated hundreds of years ago, during which time his or her family multiplied and branched out. Humans do nothing better than migrate, mix, and mate. Scientists, tending to change their countries of residence, are particularly mobile. However, names may well reflect political circumstances.

Salient features are the great frequency of British, German, Japanese, Chinese, and Indian names. One also notes a disproportionate frequency of name of Jewish and Irish origin. Some languages, mostly from Africa, are represented by a single individual, probably a graduate student who ventured into bacteriophage research. The number of names from Muslim countries, especially of Arabic countries, Pakistan, and

©2012 Landes Bioscience. Do not distribute.

Table 1. Analysis of 30 000 articles from 1965 to 2010 (14,629 author names)

Europe-related			Other		
Domain or area	**Scientists**	**%**	**Domain or area**	**Scientists**	**%**
Albanian	1		**African (sub-Saharan)**	40	
Armenian	47		• Bantu	10	
Baltic	56		• Fon	1	
British	3,112	21.5	• Ibo	4	
• English	2,257		• Yoruba	7	
• Cornish	103		• West African	8	
• Scottish	527		• Other	10	
• Welsh	225		**Amerindian**	3	
Bulgarian	17		• Quechua	2	
Czechoslovakian	173		• Siouan	1	
Dutch-Flemish	473	3.3	Amharic	9	
Finnish	90		Arabic	237	1.6
Francian	729	5.0	**Austronesian**	31	
• French	700		• Filipino	11	
• Breton	29		• Malagasy	2	
Georgian	58		• Malayan	18	
German	1,919	13.3	Berber-Kabyle	12	
Greek	92		Chinese	836	5.8
Hungarian	121		Hebrew (Israel)	69	
Irish	338		Indian	649	4.5
Iberian	719	5.0	Indo-Aryan	3	
• Spanish-Catalan	525		Iran (Farsi)	51	
• Portuguese	111		Japanese	950	6.6
• Basque	83		Korean	157	1.1
Italian	560	3.9	Mongol	10	
Jewish	609	4.2	Thai	29	
Polish	349	2.4	Turkic	45	
Romanian	56		Vietnamese	46	
Russian	711	4.9			
Scandinavian	434	3.0			
Serbo-Croatian	97		Nonidentified	232	1.6
Ukrainian-Belarus	155	1.1	Anonymous	104	0.7
Total	**10,916**	**75.6**		**3,513**	**24.3**

Indonesia, is disproportionately low when compared with the fact that Muslim countries comprise almost 25% of mankind. To the praise of small countries, one notices a high activity in phage research in Georgia, Finland, Hungary, Israel, the Netherlands, and (South) Korea.

Quirks and Oddities

Linguists may wonder that very old languages, either pre-Indoeuropean such as Basque or near-extinct Celtic languages such as Cornish, still persist in name lists. Clearly, personal names, just as those of geographical features, are inherently stable. By contrast, the dearth of names derived from Quechua and Nahua, the dominant languages of the Inca and Aztec empires, respectively, is essentially attributed to the total suppression of the Inca and Aztec cultures by the conquering Spaniards and the adoption of Spanish names by the indigenous survivors.

The presence of Mongol names do not mean that their owners are Mongols. These names, Bahadur (knight), Khan (king), and Yassa (law), went into microbiology via the Turkish language and the Mughal empire in India. "Khan" is an extremely popular family name in Pakistan. "Yassa" or "Yassak" is the law collection of Genghis Khan. Indian surnames also include that of Humayun, a Mughal emperor of the 16th century, and a number of gods. Among others, we find Lord Krishna, Lord Shiva, and Lakshmidevi, the Hindu goddess of wealth. We also find Western deity names such as Hercules and Diana, and even the Devil. However, the name which I prefer and cherish most, is El-Ghoul. A "ghoul" is a character from Arabic folklore who lives in cemeteries and eats corpses.

Conclusions

Microbiologists with British and German names predominate conspicuously. Clearly, this is related to the present dominant position of the USA and the English language in science,[5,8] but calls for more explanations. They must be sought in history and society. Assuming that all peoples are equal with respect to intelligence, what happened? I have only partial answers .

(1) Immigration to the USA is certainly one of the answers. According to a census from 2000, 37% of US citizens were of German, British, or Irish descent (15.2, 11 and 10.8%, respectively).[11] This correlates quite well with the phage statistics of **Table 1**. The British name group may be artificially inflated, including immigrants who anglicized themselves and Afro-American slaves with British names. Germans arrived in large numbers in several waves, for example after the failure of the 1848 revolution. Overall, German immigrants predominated from 1820 to 1978. People of German descent are also very common in Pennsylvania, the site of several important US universities. A wave of German scientists emigrated to the US and Australia around 1950–1955 when the outlook for German science was notoriously poor. I recall that there was no funding and that universities were paralyzed by ultra-leftists. The case of Irish scientists is particularly interesting. In the 19th century, the Irish population was in a dire situation and large numbers of people fled

©Landes Bioscience. Do not distribute.

Table 2. Complex language domains

Domain or area	Comprises
Baltic	Latvian, Lithuanian
English	Anglo-Saxon, Normannic French
Finnish	Estonian, Lappish
German	Sorbian (Wendish)
Indian	India, Sri Lanka, Nepal
Indo-Aryan	Balochi, Sindhi
Malayan	Indonesia, Malaysia
Russian	Russified Asiatic and Caucasian names
Scandinavian	Danish, Icelandic, Norvegian, Swedish
Serbo-Croatian	Slovenian
Turkic	Turkey, any Turk people
West African	Burkina Faso, Ghana, Ivory Coast, Senegal
Other African	Cameroon, Kenya, Uganda

to the US and Canada to escape poverty, famine, and oppression. By contrast, there are now hundreds of phage scientists with Irish names in North America.

(2) Historical factors are also conspicuous in the cases of India, China, Japan, and South Korea. Both India and China have become research powerhouses, expanding substantially their scientific output in the past 10 y.[12,13] The rise of Asian scientists is very evident in, for example, "Archives in Virology" and "Virological Journal." Further, in visiting India two years ago, I found that scientific bookstores carry large numbers of books on biochemistry, microbiology, and genetics, written by Indian scientists and published in India. If Indian and Chinese scientists are now prominent in phage research, this was not always so. Indeed, the first phage bibliography by Raettig[1] lists only 24 Indian and 12 Chinese first authors for the years of 1917 to 1956 (Ackermann, unpublished).

(a) The primary cause for the emergence of India must be its political independence in 1947. From then, on Indians developed universities, libraries, and research institutions.

(b) China always had a distinguished history of research and innovation. For a time, Chinese scientists almost disappeared from phage research, but are now back in force. Part of their publications may emanate from the US, Singapore or Taiwan, but most are now from Mainland China. The principal reason seems to be political and, specifically, the demise of Maoism. Mainland China started economic reforms in 1978 with the results that it has now the second-largest economy after the US. Scientists are no longer sent to reeducation, no longer have to apologize for "bourgois individualism," but receive funds and are now publishing on a grand scale.

(c) Japan was always active in phage research. Japanese researchers accounted for 6.5% of phage scientists in 1917–1957.[1] This was a consequence of the Meiji restoration in 1867 (Meiji meaning "enlightened rule" in Japanese) and the towering personality of emperor Mutsohito (1852–1912), called the Great. It changed Japan from a feudal state to a market economy and introduced Western science to Japan. Fortunately, this coincided with the emergence of microbiology in Europe. We may remember from these times the names of the great Japanese microbiologists, Kitasato and Noguchi. The historic factor is thus represented by the Government and even a single individual, the emperor of Japan.

(d) Korea was virtually absent from phage research before 1956. First subjected to Japanese occupation and badly destroyed during the Korean war, it could not engage in science. When I visited Seoul in 1985, universities were racked by student strife and nothing indicated that South Korea could be anything else than an intellectual backwater. This has changed very much. We now see a surge in the numbers of South Korean scientists. By contrast, virtually no phage publication has appeared from North Korea.

(3) Jewish scientists were always prominent in phage research (and perhaps all of virology). Vol. II of the first Raettig bibliography[1] shows that at least 6.5% of phage publications in 1917–56 originated from Jewish scientists (Ackermann H.-W., unpublished). Their contribution is now around 4% (**Table 1**), to which one must add the very important research effort of Israel. The most likely explanation for this high incidence is sociological. Throughout history, Jews always valued education and learning. I suspect that they also put pressure on their youngsters ("you, my little guy, you will pass your exams or else"). In Germany, this meant that many Jews became highly educated and part of the so-called "Bildungsbürgertum," that is "knowledge-bourgeoisie." Considering that many Jewish scientists cannot be identified by their surnames (it has been quipped that "a Jew is someone who says that he or she is a Jew"), the contribution of Jewish scientists to phage research is likely to be much higher than evident in **Table 1**.

(4) Arabic and, more generally, Muslim countries participate little in phage research. Moreover, a number of scientists with Arab or Iranian names are expatriates and work (or worked) in the West. This has no religious connotations as scientists with Arabic names may be Copts or Lebanese Christians. As reported in a recent, detailed study on medical research,[9] the most productive Islamic countries are Turkey and Iran, whereas large, populous countries such as Pakistan and Indonesia are essentially inactive in research. Given the great historical role of Islamic scientists in preserving and advancing science during the Middle Ages and the deep respect for knowledge that I saw myself in Islamic countries, this is quite incomprehensible. How did this happen? A superficial investigation on the Net shows (a) frequent referral to the past, Scripture and religious authorities, and (b) very critical papers, all by Islamic authors, which recuse the role of religion in science.[14-16] This debate underlines how much science depends on humans.

(5) It is heartening so see that people from various African countries have now

©2012 Landes Bioscience. Do not distribute.

moved into phage research. One can only hope that this continues. Phage research is not only big-ticket science linked to expensive instruments (e.g., electron microscopes) or DNA sequencing. Indeed, much is to be gained by investigating relatively inexpensive subjects like ecology and phage therapy.

(6) Ultimately, the relative frequency of Irish and Jewish scientists and the rise of phage research in Japan, China and India, all due to historic and societal circumstances, are difficult to explain by the structural and numerical theories on the advancement of science,[6] but are in agreement with or confirm Merten's cultural theory.[6]

Acknowledgments

Linguistic help was provided by Dr. Anastasia Aksyuk, Purdue University, West Lafayette, IN; Zahira Chouchane, Laval University, Quebec; Vidya Desai, GangaGen, Bangalore, India; Dr. Alla Kushkina, Institute of Microbiology, Kyiv, Ukraine; and Dr. Ipek Kurtböke, Sunshine Coast University, Qld., Australia, for help in identifiying Arabic, Berber, Indian, Turkish and Ukrainian names. Ms. Vivianne Gélinas, Quebec, verified the English style of this text.

References

1. Raettig H. Bakteriophagie 1917 bis 1956. Vol. II. Autorenregister. Gustav Fischer 1958.
2. Bakteriophagie RH. 1957-1965. Vol II. Bibliographie. Gustav Fischer 1967.
3. Ackermann H-W. Obituary: Hansjürgen Raettig - Collector of bacteriophage references. BEG News 3:5, http://www.phage.org/, 2000.
4. Abedon ST. Bacteriophage Ecology Group (BEG) Bibliography, http://www.mansfield.ohio-state.edu/~sabedon/beg_bibliography.htm, 2007.
5. Garfield E. Mapping the world of science. 150th Anniv Amer Assoc Advan Sci, Philadelphia, PA, http://garfield.library.upenn.edu/oaoers/mapssciworld.html, 1998.
6. Merton RK. Science, Technology and Society in Seventeenth Century England. Howard Fertig Press (reissued) 1970. First published in 1938 by St. Catherine Press.
7. Cole S, Phelan TJ. The scientific productivity of nations. Minerva 1999; 37:1-23; http://dx.doi.org/10.1023/A:1004591413630
8. King C. Which countries publish the most cited scientific papers? Science Watch 2007. Also in Basu A. Scientific productivity and citation of scientific papers: where do we stand? Curr Sci 2007; 93:750-1.
9. Habibi E, Mirhosseini Z, Majidi M. Medical publications (2002–2009) of Islamic countries: a Medline-based study compared to non-Islamic countries. Iran J Med Sci 2010; 35:226-35.
10. Kiljunen S, Datta N, Dentovskaya SV, Anisimov AP, Knirel YuA, Bengoechea JA, et al. Identification of the lipopolysaccharide core of *Yersinia pestis* and *Yersinia pseudotuberculosis* as the receptor for bacteriophage φA1122. J Bacteriol 2011; 193:4963-72; PMID: 21764935; http://dx.doi.org/10.1128/JB.00339-11
11. Anonymous. U.S. Census Bureau, Summary File 3. http//enwikipedia.org, 2000
12. Adams J, King C, Ma N. The new geography of science: Research and collaboration in China. Global Resarch Reports November 2009, http://researchanalytics.thompsonreuters.com/grr, 2009.
13. Adams J, King C, Singh V. The new geography of science. Research and collaboration in India. Global Research Reports October 2009, http://researchanalytics.thomsonreuters.com/grr, 2009.
14. Anonymous. Fundamental extremist Islamic Pakistan: A huge suicide-bomb and its own worst enemy, http://www.secularsites.freeuk.com.science_and_islam.htm, 2004.
15. Farooqi MIH. Status of Muslim societies, www.parvez-videocom/muslim_societis.asp, 2004.
16. Mahmud I. Science, technology and the Muslim world - from a Golden Age to stagnation. Khan Bahadur Ahsanullah Memorial Lecture, www.aust.edu/publication/khaml1.doc, 2009.

© 2012 Landes Bioscience.
Do not distribute.